

**0 7 2 6 3 6 0**

(i) APPLICANT: Neuhold, Lisa  
Killar, Loran

(iii) NUMBER OF SEQUENCES: 21

(A) ADDRESSEE: Darby & Darby PC  
(B) STREET: 805 Third Avenue  
(C) CITY: New York  
(D) STATE: NY  
(E) COUNTRY: USA  
(F) ZIP: 10022

(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(A) APPLICATION NUMBER: 08/994,689  
(B) FILING DATE: 1997-12-19  
(C) CLASSIFICATION:

(A) APPLICATION NUMBER:  
(B) FILING DATE:

(A) NAME: Green, Reza  
(B) REGISTRATION NUMBER: 38,475  
(C) REFERENCE/DOCKET NUMBER: 0630/0D532

(A) TELEPHONE: 212-527-7700  
(B) TELEFAX: 212-753-6237  
(C) TELEX: 236687

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	His	Pro	Gly	Val	Leu	Ala	Ala	Phe	Leu	Phe	Leu	Ser	Trp	Thr	His		
1				5				10						15			
Cys	Arg	Ala	Leu	Pro	Leu	Pro	Ser	Gly	Gly	Asp	Glu	Asp	Asp	Leu	Ser		
			20					25					30				
Glu	Glu	Asp	Leu	Gln	Phe	Ala	Glu	Arg	Tyr	Leu	Arg	Ser	Tyr	Tyr	His		
		35					40					45					
Pro	Thr	Asn	Leu	Ala	Gly	Ile	Leu	Lys	Glu	Asn	Ala	Ala	Ser	Ser	Met		
	50				55						60						
Thr	Glu	Arg	Leu	Arg	Glu	Met	Gln	Ser	Phe	Phe	Gly	Leu	Glu	Val	Thr		
65					70					75				80			
Gly	Lys	Leu	Asp	Asp	Asn	Thr	Leu	Asp	Val	Met	Lys	Lys	Pro	Arg	Cys		
			85					90						95			
Gly	Val	Val	Asp	Val	Gly	Glu	Tyr	Asn	Val	Phe	Pro	Arg	Thr	Leu	Lys		
			100					105					110				
Trp	Ser	Lys	Met	Asn	Leu	Thr	Tyr	Arg	Ile	Val	Asn	Tyr	Thr	Pro	Asp		
		115					120					125					
Met	Thr	His	Ser	Glu	Val	Glu	Lys	Ala	Phe	Lys	Lys	Ala	Phe	Lys	Val		
	130					135					140						
Trp	Ser	Asp	Val	Thr	Pro	Leu	Asn	Phe	Thr	Arg	Leu	His	Asp	Gly	Ile		
145					150					155					160		
Ala	Asp	Ile	Met	Ile	Ser	Phe	Gly	Ile	Lys	Glu	His	Gly	Asp	Phe	Tyr		
			165					170					175				
Pro	Phe	Asp	Gly	Pro	Ser	Gly	Leu	Leu	Ala	His	Ala	Phe	Pro	Pro	Gly		
			180				185						190				
Pro	Asn	Tyr	Gly	Gly	Asp	Ala	His	Phe	Asp	Asp	Asp	Glu	Thr	Trp	Thr		
	195				200							205					
Ser	Ser	Ser	Lys	Gly	Tyr	Asn	Leu	Phe	Leu	Val	Ala	Ala	His	Glu	Phe		
	210				215						220						
Gly	His	Ser	Leu	Gly	Leu	Asp	His	Ser	Lys	Asp	Pro	Gly	Ala	Leu	Met		
225					230					235					240		
Phe	Pro	Ile	Tyr	Thr	Tyr	Thr	Gly	Lys	Ser	His	Phe	Met	Leu	Pro	Asp		
			245					250					255				
Asp	Asp	Val	Gln	Gly	Ile	Gln	Ser	Leu	Tyr	Gly	Pro	Gly	Asp	Glu	Asp		
			260					265					270				
Pro	Asn	Pro	Lys	His	Pro	Lys	Thr	Pro	Asp	Lys	Cys	Asp	Pro	Ser	Leu		
	275					280						285					
Ser	Leu	Asp	Ala	Ile	Thr	Ser	Leu	Arg	Gly	Glu	Thr	Met	Ile	Phe	Lys		
	290				295						300						
Asp	Arg	Phe	Phe	Trp	Arg	Leu	His	Pro	Gln	Gln	Val	Asp	Ala	Glu	Leu		
305					310				315						320		
Phe	Leu	Thr	Lys	Ser	Phe	Trp	Pro	Glu	Leu	Pro	Asn	Arg	Ile	Asp	Ala		
			325					330					335				

Ala Tyr Glu His Pro Ser His Asp Leu Ile Phe Ile Phe Arg Gly Arg  
 340 345 350  
 Lys Phe Trp Ala Leu Asn Gly Tyr Asp Ile Leu Glu Gly Tyr Pro Lys  
 355 360 365  
 Lys Ile Ser Glu Leu Gly Leu Pro Lys Glu Val Lys Lys Ile Ser Ala  
 370 375 380  
 Ala Val His Phe Glu Asp Thr Gly Lys Thr Leu Leu Phe Ser Gly Asn  
 385 390 395 400  
 Gln Val Trp Arg Tyr Asp Asp Thr Asn His Ile Met Asp Lys Asp Tyr  
 405 410 415  
 Pro Arg Leu Ile Glu Glu Asp Phe Pro Gly Ile Gly Asp Lys Val Asp  
 420 425 430  
 Ala Val Tyr Glu Lys Asn Gly Tyr Ile Tyr Phe Phe Asn Gly Pro Ile  
 435 440 445  
 Gln Phe Glu Tyr Ser Ile Trp Ser Asn Arg Ile Val Arg Val Met Pro  
 450 455 460  
 Ala Asn Ser Ile Leu Trp Cys  
 465 470

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 470 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	60
TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	TTTACCACTC	CCTATCAGTG	ATAGAGAAAA	120
GTGAAAGTCG	AGTTTACCAC	TCCCTATCAG	TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	180
ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	240
AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	300
CTCGGTACCC	GGGTCGAGTA	GGCGTGTACG	GTGGGAGGCC	TATATAAGCA	GAGCTCGTTT	360
AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAGACA	420
CCGGGACCGA	TCCAGCCTCC	GCGGCCCCGA	ATTAGCTTGA	TATCGAATTC		470

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3479 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGTACCACTA GTAAGCTTAG ATCCACTGTC TGGGATTATA TCAGGACAAC CGAAGCCTGG 60

AAAGTGTATT	AGGTAGAGCA	TTTTCTTCCA	CGTGTTTGGG	CACGTTTCCG	ACAGCTAGGA	120
TTCCAGCTCT	GTCTTTGTAT	GTTACAGACT	GTAAATCAAT	CGCAGGTGAA	ACTGTTTGGA	180
CAGTAGGTGG	GGATCAAAGA	CCCTCCGCCC	GTGAGACTCT	AGGCGCTTTC	CCCTGCCACC	240
AGCCTGTCTC	CAGAGATGCT	CTGGAAGGAG	GCGGGCCCCG	GCGGTCTTTC	TGCTCTTTAG	300
CGTGGCGGAC	GCGGCGGCGG	GGGCAGGGCT	GGAGCAGAGA	GCGCTGCAGT	GATAGAACTT	360
TCTGACCCCG	CTGCGCAGGG	CGGCAGGGTG	GCAGGGTGGC	AGGGTGGCGA	GCTAAGCCAG	420
AGCCGAACGC	TGGAGCTCTG	GGAGGAACAT	CGAAGGTTTG	TATGTGGTCT	GAGATCGGCC	480
TGACTATATT	TTTTTGTCTT	AAATTTGCAA	GCACACACCC	ACAAAGCTGC	GGTCTTGACC	540
GGTATTCTTT	ATAGAGCGCA	ATGGAGTGAG	CTGAGTGTCT	AAACGATTTC	CCTAATTCAT	600
CTGATAGCAG	AGGCGCTCTC	CTAATTGGCG	AAGAGCTGCC	TCATGTCCGC	AACTTTTTTG	660
CAGAGTGAAT	TCCACAGCTT	TGTGTGTGTG	TGTGGGGGGG	GGTGTAAAGG	GTGTCTAAAA	720
CTTTCGGTCT	CCTACTATTC	TGTATCTCGA	CCGGTTGGTT	TTACACCCCG	GCTCATCTCA	780
TCAACGCAAA	CACCCCCACT	CTCCTATGGA	CCCAAGGACC	TGACGTGGGG	GAAGGTGGAC	840
ATTAGGAATG	TCAGAAACCT	AGAGTCCACG	CTCCTCCTCT	CCATCTTTCC	ACGAGTTTGG	900
GAAACTTCTT	GGCTGCGAAG	ACTTTGACCC	ACATCTGCAT	TTCTCAGCCC	CAGCTTCCAA	960
AAGTGCTGCA	GGTTCGGGAG	GGGAGACCTC	AGTCCTCCTT	TGTGAGGCTT	GTTTGCGTTG	1020
AGGGATTGGC	AGCGATGGCT	TCCAGATGGG	CTGAAACCCCT	GCCCGTATTT	ATTTAAACTG	1080
GTTCCCTCGTG	GAGAGCTGTG	AATCGGGCTC	TGTATGCGCT	CGAGAAAAGC	CCCATTCTATG	1140
AGAGGCAAGG	CCCACTGGGT	CCCCCCGACT	CCCCGACCCC	CCTCTCCCAC	AATATATCCC	1200
CCCTCCCTGT	GCCCGCCTGC	CGCCACCTCC	CGGGCTCCGG	CCCCGCGCGC	AGCGGCGACG	1260
AAGCAACACA	GTTCCCCGAA	AGAGGTAGCT	TTTTAATTGG	CCAGCCACAA	AGAATCACTT	1320
ATGCCGCACG	GCGGTAACGA	GGGGAACCGG	ATCGGGCGGC	CAGGATGCTA	TCTGTGTAGC	1380
CCTTTTCGTG	CCACAATTAG	GGTGGTGCTG	GCTTCCTCCG	ACCGCACCTA	GGCGATCTGG	1440
TTACACTGTT	GGCTCCTTTC	TTGGGCAGTC	ATTTAATCCT	ACTTTTTACT	CTACGAATGT	1500
CTGTCTGATG	GAGGGCTGTG	TCCGGAGCCC	CATCCACAAA	GAGTCAGCCA	GCAGCTCTCA	1560
CACCCGGCTG	GATCTCATAT	GGTGCACCTC	CAGTACAATC	TGCTCTGATG	CCGCATAGTT	1620
AAGCCAGCCA	AGCTAGCTTG	CGCAAGCTAG	CTTGCGATCC	GTAAAAATGT	GTGAGAGTTA	1680
CAAAATGTCT	TCCGGGCTAA	GATCCGACAG	CCATGGTCCA	AAGAAGACTT	CGGCACTGCA	1740
GACTTAAAAC	CAGCTTTTCTA	GCAGAGGCAG	AAGGATCTAG	AGCCAAAGGC	AAAGACTTGA	1800
ATAGGCTGGG	AAGATGCAAG	AATGGCATT	TACATAAAGA	ACACTCTCTC	CTTTTCCAGC	1860
CAGCACACTT	GCATAGAAAT	TAAGTTTAC	ACTTGAAGTT	CTTTGTTTCC	ATCCTGAGAA	1920
GCTCCAAAGT	CTGAGGTGGT	GTGGTATGCT	GGGTAATTCT	CCCCACCCCC	CAACATTCCC	1980
TGGGGGTTCC	ATGGGGGTAG	CTTCTCCCAA	GGACTTCCAG	CGGCAACACA	GAAATCCCAC	2040
TTTCGAGACAA	AGGAGTTACT	GCTTAAATCA	GGCCCTAATT	TCCAAGGTTT	CCTTTGCTTA	2100
AAGTTCCCTA	GAGGACCATC	TCACTTCTAA	AGAAAAGGTG	TATTCGGGGA	CCCATCCTCA	2160
ACCTCCTTGT	TATGGAAGGA	GACTTCGGGA	ACAGAGCAAG	GGCTGAGCCT	CCGGCAGTTT	2220
GGGGTAAGGT	TGGGGTTGGG	GGGAGCAAGG	AAGGCAAGTG	AGGCTGGAGG	CCCAGGGATA	2280
GGGGAAGATG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTC	TCGGGGATGG	2340
TGGTGGTGGA	CAACTAGGAA	ACTCTGGCGC	TTTCTCCTCC	CCTCACAAAA	CTGAGTCCAG	2400
CTGGAGCCGC	CTCCAGACTC	TCTGGCCAGG	GCCTCAGAGT	GGTCAACAGT	CCCTGGCCAG	2460
CGTTGCTCTC	TCCAGGCTAA	GGGCACCCAC	TCCCCTGGAG	ATTCTTGAAC	CTGGGCCAGG	2520
AAGAGCCGAA	TTAGACAAGT	GTCTCCAATC	CGGCTGCGTG	CGGATTTTGT	TGCGGTGTCC	2580
CTCGGTTGTC	TGCAGTTCCT	TTAGTCCCTT	CCCTGGCCTG	CCCCTTACAC	CTCCACACAG	2640
GTCCCCCTCT	GTGTAGGAAT	ACACCAGACC	CTCTCTTAGC	CACACACACC	TCCAGTCCCC	2700
CGTCTACCTA	GATTTTTTTC	ATAGCTAGTT	GGATGGGGGA	TGGGTTAGGG	AGGCTGGGTT	2760
TGCGAGCCTC	CAGGTGGGAG	TTCACCGACA	GGTACTCCGC	AAAGGAGCTG	GAAGGCAGGT	2820
CTGGAAAACT	GTCCCCCAGA	TTTAGGATT	TGGGCAGCTT	CCATCAGCTT	ATACTTTGGC	2880
TCCCCCGCCC	CCTAAACTCC	CCATCCCCAC	CTTCCTTTCT	CCCGTTACTT	CGTCCTCCCT	2940
CGCCTTTCCA	GCCTTGAGTC	TAAAGCTCCA	TGCTTATGCC	TCTGCAAAAC	ACCCCTCCC	3000
TTCTAACCCC	AGCAGAACTC	CGAGGAAAGG	GGCCGGAGGC	CCCCCTTCTC	GCCTGTGGTT	3060
AGAGGGGGCA	GTGTGGCAGT	CCCAAGTGGG	GGCGACCGGA	GGCCGTCTCG	GTGCCCCGCC	3120

CGATCAGGCC	ACTGGGCACA	TCGGGGGCGG	GAAGCTGGGC	TCACCAAAGG	GGCGACTGGC	3180
CTTGGCAGGT	GTGGGCTCTG	GTCCGGCCTG	GGCAGGCTCC	GGGGGCGGGG	TCTCAGGTTA	3240
CAGCCCCGCG	GGGGGCTGGG	GGGCGGCCCC	CGGTTTGGGC	TGGTTTGCCA	GCCTTTGGAG	3300
CGACCGGGAG	CATATAACCG	GAGCCTCTGC	TGGGAGAAGA	CGCAGAGCGC	CGCTGGGCTG	3360
CCGGGTCTCC	TGCCTCCTCC	TCCTGCTCCT	AGAGCCTCCT	GCATGAGGGC	GCGGTAGAGA	3420
CCCGGACCCG	CTCCGTGCTC	TGCCGCCTCG	CCGAGCTTCG	CCCGCAAGCT	GGGGAATTC	3479

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Pro Arg Cys Gly Val Pro Asp Val  
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAAGCCAAGAT	GCGGGGTTGT	CGATGTGGGT	GAATACAAT	39
-------------	------------	------------	-----------	----

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GAAAAAGCCA	AGATGCGGGG	GTCCTGATGT	GGGTGAATAC	40
------------	------------	------------	------------	----

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGTACCACTA	GTAAGCTTAG	ATCTCATATG	GTCGACCCCG	GGGAATTCCT	GCAGGGATCC	60
TCTAGAAGTA	CTCCATGGGT	ATACATCGAT	GCGGCCCG			98

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2792 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	60
TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	TTTACCACTC	CCTATCAGTG	ATAGAGAAAA	120
GTGAAAGTCG	AGTTTACCAC	TCCCTATCAG	TGATAGAGAA	AAGTGAAAGT	CGAGTTTACC	180
ACTCCCTATC	AGTGATAGAG	AAAAGTGAAA	GTCGAGTTTA	CCACTCCCTA	TCAGTGATAG	240
AGAAAAGTGA	AAGTCGAGTT	TACCACTCCC	TATCAGTGAT	AGAGAAAAGT	GAAAGTCGAG	300
CTCGGTACCC	GGGTCGAGTA	GGCGTGTACG	GTGGGAGGCC	TATATAAGCA	GAGCTCGTTT	360
AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAGACA	420
CCGGGACCGA	TCCAGCCTCC	GCGGCCCCGA	ATTAGCTTGA	TATCGAATTC	GAGCTCGGTA	480
CCCGGGGATC	CTCTAGACAA	GATGCATCCA	GGGGTCTTGG	CTGCCTTCCT	CTTCTTGAGC	540
TGGACTCATT	GTCGGGCCCT	GCCCCTTCCC	AGTGGTGGTG	ATGAAGATGA	TTTGTCTGAG	600
GAAGACCTCC	AGTTTGCAGA	GCGCTACCTG	AGATCATACT	ACCATCCTAC	AAATCTCGCG	660
GGAATCCTGA	AGGAGAATGC	AGCAAGCTCC	ATGACTGAGA	GGCTCCGAGA	AATGCAGTCT	720
TTCTTCCGGCT	TAGAGGTGAC	TGGCAAACCTT	GACGATAACA	CCTTAGATGT	CATGAAAAAG	780
CCAAGATGCG	GGGTTGTCTGA	TGTGGGTGAA	TACAATGTTT	TCCCTCGAAC	TCTTAAATGG	840
TCCAAAATGA	ATTTAACCTA	CAGAATTGTG	AATTACACCC	CTGATATGAC	TCATCTTGAA	900
GTCGAAAAGG	CATTCAAAAA	AGCCTTCAAA	GTTTGGTCCG	ATGTAACCTC	TCTGAATTTT	960
ACCAGACTTC	ACGATGGCAT	TGCTGACATC	ATGATCTCTT	TTGGAATTAA	GGAGCATGGC	1020
GACTTCTACC	CATTTGATGG	GCCCTCTGGC	CTGCTGGCTC	ATGCTTTTCC	TCCTGGGCCA	1080
AATTATGGAG	GAGATGCCCA	TTTTGATGAT	GATGAAACCT	GGACAAGTAG	TTCCAAAGGC	1140
TACAACTTGT	TTCTTGTTCG	TGCGCATGAG	TTCGGCCACT	CCTTAGGTCT	TGACCACTCC	1200
AAGGACCTTG	GAGCACTCAT	GTTTCCTATC	TACACCTACA	CCGGCAAAAG	CCACTTTATG	1260
CTTCCTGATG	ACGATGTACA	AGGGATCCAG	TCTCTCTATG	GTCCAGGAGA	TGAAGACCCC	1320
AACCCTAAAC	ATCCAAAAAC	GCCAGACAAA	TGTGACCCTT	CCTTATCCCT	TGATGCCATT	1380
ACCAGTCTCC	GAGGAGAAAC	AATGATCTTT	AAAGACAGAT	TCTTCTGGCG	CCTGCATCCT	1440
CAGCAGGTTG	ATGCGGAGCT	GTTTTTAACG	AAATCATTTT	GGCCAGAACT	TCCCAACCGT	1500
ATTGATGCTG	CATATGAGCA	CCCTTCTCAT	GACCTCATCT	TCATCTTCAG	AGGTAGAAAA	1560
TTTTGGGCTC	TTAATGGTTA	TGACATTCTG	GAAGGTTATC	CCAAAAAAT	ATCTGAACTG	1620
GGTCTTCCAA	AAGAAGTTAA	GAAGATAAGT	GCAGCTGTTC	ACTTTGAGGA	TACAGGCAAG	1680

ACTCTCCTGT	TCTCAGGAAA	CCAGGTCTGG	AGATATGATG	ATACTAACCA	TATTATGGAT	1740
AAAGACTATC	CGAGACTAAT	AGAAGAAGAC	TCCCCAGGAA	TTGGTGATAA	AGTAGATGCT	1800
GTCTATGAGA	AAAATGGTTA	TATCTATTTT	TTCAACGGAC	CCATACAGTT	TGAATACAGC	1860
ATCTGGAGTA	ACCGTATTGT	TCGCGTCATG	CCAGCAAATT	CCATTTTGTG	GTGTTAAGTG	1920
TCTTTTTTAAA	AATTGTTATT	TAAATCCTGA	AGAGCATTTG	GGGTAATACT	TCCAGAAGTG	1980
CGGGGTAGGG	GAAGAAGAGC	TATCAGGAGA	AAGCTCTAGT	TCTAGAGGGC	CCTATTCTAT	2040
AGTGTCACCT	AAATGCTAGA	GGATCTTTGT	GAAGGAACCT	TACTTCTGTG	GTGTGACATA	2100
ATTGGACAAA	CTACCTACAG	AGATTTTAAAG	CTCTAAGGTA	AATATAAAAT	TTTTAAGTGT	2160
ATAATGTGTT	AAACTACTGA	TTCTAATTGT	TTGTGTATTT	TAGATTCCAA	CCTATGGAAC	2220
TGATGAATGG	GAGCAGTGGT	GGAATGCCTT	TAATGAGGAA	AACCTGTTTT	GCTCAGAAGA	2280
AATGCCATCT	AGTGATGATG	AGGCTACTGC	TGACTCTCAA	CATTCTACTC	CTCCAAAAAA	2340
GAAGAGAAAG	GTAGAAGACC	CCAAGGACTT	TCCTTCAGAA	TTGCTAAGTT	TTTTGAGTCA	2400
TGCTGTGTTT	AGTAATAGAA	CTCTTGCTTG	CTTTGCTATT	TACACCACAA	AGGAAAAAGC	2460
TGCACTGCTA	TACAAGAAAA	TTATGGAAAA	ATATTTGATG	TATAGTGCCT	TGACTAGAGA	2520
TCATAATCAG	CCATACCACA	TTTGTAGAGG	TTTTACTTGC	TTTAAAAAAC	CTCCACACC	2580
TCCCCCTGAA	CCTGAAACAT	AAAATGAATG	CAATTGTTGT	TGTTAACTTG	TTTATTGCAG	2640
CTTATAATGG	TTACAAATAA	AGCAATAGCA	TCACAAATTT	CACAAATAAA	GCATTTTTTT	2700
CACTGCATTC	TAGTTGTGGT	TTGTCCAAAC	TCATCAATGT	ATCTTATCAT	GTCTGGATCA	2760
TCCCGCCATG	GGTATACATC	GATGCGGCCG	CC			2792

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5276 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGTACCACTA	GTAAGCTTAG	ATCCACTGTC	TGGGATTATA	TCAGGACAAC	CGAAGCCTGG	60
AAAGTGATTT	AGGTAGAGCA	TTTTCTTCCA	CGTGTTTGGG	CACGTTTCCG	ACAGCTAGGA	120
TTCCAGCTCT	GTCTTTGTAT	GTTACAGACT	GTAAATCAAT	CGCAGGTGAA	ACTGTTTGGA	180
CAGTAGGTGG	GGATCAAAGA	CCCTCCGCCC	GTGAGACTCT	AGGCGCTTTC	CCCTGCCACC	240
AGCCTGTCTC	CAGAGATGCT	CTGGAAGGAG	GCGGGCCCGG	GCGGTCTTTC	TGCTCTTTAG	300
CGTGGCGGAC	GCGGCGGCGG	GGGCAGGGCT	GGAGCAGAGA	GCGCTGCAGT	GATAGAACTT	360
TCTGACCCCG	CTGCGCAGGG	CGGCAGGGTG	GCAGGGTGGC	AGGGTGCGGA	GCTAAGCCAG	420
AGCCGAACGC	TGGAGCTCTG	GGAGGAACAT	CGAAGTGTTT	GTATGTGGTC	TGAGATCGGC	480
CTGACTATAT	TTTTTTGTCC	TAAATTTGCA	AGCACACACC	CACAAAGCTG	CGGTCTTGAC	540
CGGTATTCTT	TATAGAGCGC	AATGGAGTGA	GCTGAGTGTC	TAAACGATTT	CCCTAATTCA	600
TCTGATAGCA	GAGGCGCTCT	CCTAATTGGC	GAAGAGCTGC	CTCATGTCCG	CAACTTTTTG	660
GCAGAGTGAA	TTCCACAGCT	TTGTGTGTGT	GTGTGGGGGG	GGGTGTAAGG	GGTGTCTAAA	720
ACTTTCGGTC	TCCTACTATT	CTGTATCTCG	ACCGGTTGGT	TTTACACCCC	GGCTCATCTC	780
ATCAACGCAA	ACACCCCCAC	TCTCCTATGG	ACCCAAGGAC	CTGACGTGGG	GGAAGGTGGA	840
CATTAGGAAT	GTCAGAAACC	TAGAGTCCAC	GCTCCTCCTC	TCCATCTTTC	CACGAGTTTG	900
GGAAACTTCT	TGGCTGCGAA	GACTTTGACC	CACATCTGCA	TTTCTCAGCC	CCAGCTTCCA	960
AAAGTGCTGC	AGGTTGCGGA	GGGGAGACCT	CAGTCCTCCT	TTGTGAGGCT	TGTTTGCGTT	1020
GAGGGATTGG	CAGCGATGGC	TTCCAGATGG	GCTGAAACCC	TGCCCCGATT	TATTTAAACT	1080
GGTTCCTCGT	GGAGAGCTGT	GAATCGGGCT	CTGTATGCGC	TCGAGAAAAG	CCCCATTTCAT	1140

GAGAGGCAAG	GCCCCAGTGGG	TCCCCCCGAC	TCCCCGACCC	CCCTCTCCCA	CAATATATCC	1200
CCCCTCCCTG	TGCCCCGCCTG	CCGCCACCTC	CCGGGCTCCG	GCCCCGCGCG	CAGCGGCGAC	1260
GAAGCAACAC	AGTTCCCCGA	AAGAGGTAGC	TTTTTAATTG	GCCAGCCACA	AAGAATCACT	1320
TATGCCGCAC	GGCGGTAACG	AGGGGAACCG	GATCGGGCGG	CCAGGATGCT	ATCTGTGTAG	1380
CCCTTTTCGT	GCCACAATTA	GGGTGGTGCT	GGCTTCCTCC	GACCGCACCT	AGGCGATCTG	1440
GTTACACTGT	TGGCTCCTTT	CTTGGGCAGT	CATTTAATCC	TACTTTTTTAC	TCTACGAATG	1500
TCTGTCTGAT	GGAGGGCTGT	GTCCGGAGCC	CCATCCACAA	AGAGTCAGCC	AGCAGCTCTC	1560
ACACCCGGCT	GGATCTCATA	TGGTGCATC	TCAGTACAAT	CTGCTCTGAT	GCCGCATAGT	1620
TAAGCCAGCC	AAGCTAGCTT	GCGCAAGCTA	GCTTGCGATC	CGTAAAAATG	TGTGAGAGTT	1680
ACAAAATGTC	TTCCGGGCTA	AGATCCGACA	GCCATGGTCC	AAAGAAGACT	TCGGCACTGC	1740
AGACTTAAAA	CCAGCTTTCT	AGCAGAGGCA	GAAGGATCTA	GAGCCAAAGG	CAAAGACTTG	1800
AATAGGCTGG	GAAGATGCAA	GAATGGCATT	TTACATAAAG	AACACTCTCT	CCTTTTCCAG	1860
CCAGCACACT	TGCATAGAAA	TTAAGTTTTA	CAC TTGAAGT	TCTTTGTTTC	CATCCTGAGA	1920
AGCTCCAAAG	TCTGAGGTGG	TGTGGTATGC	TGGGTAATTC	TCCCCACCCC	CCAACATTCC	1980
CTGGGGGTTC	CATGGGGGTA	GCTTCTCCCA	AGGACTTCCA	GCGGCAACAC	AGAAATCCCA	2040
CTTCGAGACA	AAGGAGTTAC	TGCTTAAATC	AGGCCCTAAT	TTCCAAGGTT	CCCTTTGCTT	2100
AAAGTTCCCT	AGAGGACCAT	CTCACTTCTA	AAGAAAAGGT	GTATTCGGGG	ACCCATCCTC	2160
AACCTCCTTG	TTATGGAAGG	AGACTTCGGG	AACAGAGCAA	GGGCTGAGCC	TCCGGCAGTT	2220
TGGGGTAAGG	TTGGGGTTGG	GGGGAGCAAG	GAAGGCAAGT	GAGGCTGGAG	GCCCAGGGAT	2280
AGGGGAAGAT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	CTCGGGGATG	2340
GTGGTGGTGG	ACAAC TAGGA	AACTCTGGCG	CTTTCTCCTC	CCCTCACAAA	ACTGAGTCCA	2400
GCTGGAGCCG	CCTCCAGACT	CTCTGGCCAG	GGCCTCAGAG	TGGTCAACAG	TCCCTGGCCA	2460
GCGTTGCTCT	CTCCAGGCTA	AGGGCACCCA	CTCCCCTGGA	GATTCTTGAA	CCTGGGCCAG	2520
GAAGAGCCGA	ATTAGACAAG	TGTCTCCAAT	CCGGCTGCGT	GCGGATTTTG	TTGCGGTGTC	2580
ECTCGGTTGT	CTGCAGTTCC	TTTAGTCCCT	TCCCTGGCCT	GCCCCTTACA	CCTCCACACA	2640
GGTCCCCCTC	TGTGTAGGAA	TACACCAGAC	CCTCTCTTAG	CCACACACAC	CTCCAGTCCC	2700
CCGTCTACCT	AGATTTTTTTT	CATAGCTAGT	TGGATGGGGG	ATGGGTTAGG	GAGGCTGGGT	2760
TTGCGAGCCT	CCAGGTGGGA	GTTCACCGAC	AGGTACTCCG	CAAAGGAGCT	GGAAGGCAGG	2820
TCTGGAAAAC	TGTCCCCCAG	ATTTAGGATT	CTGGGCAGCT	TCCATCAGCT	TATACTTTGG	2880
CTCCCCCGCC	CCCTAAACTC	CCCATCCCCA	CCTTCCTTTC	TCCCGTTACT	TCGTCTCTCC	2940
TCGCCTTTCC	AGCCTTGAGT	CTAAAGCTCC	ATGCTTATGC	CTCTGCAAAC	AACCCCTCC	3000
ETTCTAACCC	CAGCAGAACT	CCGAGGAAAG	GGGCCGGAGG	CCCCCTTCT	CGCCTGTGGT	3060
TAGAGGGGGC	AGTGTGGCAG	TCCCAAGTGG	GGGCGACCGG	AGGCCGTCTC	GGTGCCCCGC	3120
CCGATCAGGC	CAC TGGGCAC	ATCGGGGGCG	GGAAGCTGGG	CTCACCAAAG	GGGCGACTGG	3180
CCTTGGCAGG	TGTGGGCTCT	GGTCCGGCCT	GGGCAGGCTC	CGGGGGCGGG	GTCTCAGGTT	3240
ACAGCCCCGC	GGGGGGCTGG	GGGGCGGCCC	GCGGTTTGGG	CTGGTTTGCC	AGCCTTTGGA	3300
GCGACCGGGA	GCATATAACC	GGAGCCTCTG	CTGGGAGAAG	ACGCAGAGCG	CCGCTGGGCT	3360
GCCGGGTCTC	CTGCCTCCTC	CTCCTGCTCC	TAGAGCCTCC	TGCATGAGGG	CGCGGTAGAG	3420
ACCCGGACCC	GCTCCGTGCT	CTGCCGCCTC	GCCGAGCTTC	GCCCCGAAGC	TGGGGAATTC	3480
ATATGTCTAG	ATTAGATAAA	AGTAAAGTGA	TTAACAGCGC	ATTAGAGCTG	CTTAATGAGG	3540
TCGGAATCGA	AGGTTTAACA	ACCCGTAAAC	TCGCCCAGAA	GCTAGGTGTA	GAGCAGCCTA	3600
CATTGTATTG	GCATGTAAAA	AATAAGCGGG	CTTTGCTCGA	CGCCTTAGCC	ATTGAGATGT	3660
TAGATAGGCA	CCATACTCAC	TTTTGCCCCTT	TAGAAGGGGA	AAGCTGGCAA	GATTTTTTAC	3720
GTAATAACGC	TAAAAGTTTT	AGATGTGCTT	TACTAAGTCA	TCGCGATGGA	GCAAAAGTAC	3780
ATTTAGGTAC	ACGGCCTACA	GAAAAACAGT	ATGAAACTCT	CGAAAATCAA	TTAGCCTTTT	3840
TATGCCAACA	AGGTTTTTCA	CTAGAGAATG	CATTATATGC	ACTCAGCGCT	GTGGGGCATT	3900
TTACTTTAGG	TTGCGTATTG	GAAGATCAAG	AGCATCAAGT	CGCTAAAGAA	GAAAGGGAAA	3960
CACCTACTAC	TGATAGTATG	CCGCCATTAT	TACGACAAGC	TATCGAATTA	TTTGATCACC	4020
AAGGTGCAGA	GCCAGCCTTC	TTATTTCGGC	TTGAATTGAT	CATATGCGGA	TTAGAAAAAC	4080
AACTTAAATG	TGAAAGTGGG	TCCGCGTACA	GCCGCGCGCG	TACGAAAAAC	AATTACGGGT	4140
CTACCATCGA	GGGCCTGCTC	GATCTCCCGG	ACGACGACGC	CCCCGAAGAG	GCGGGGCTGG	4200



CGGCTCCGCG	CCTGTCCTTT	CTCCCCGCGG	GACACACGCG	CAGACTGTCTG	ACGGCCCCCC	4260
CGACCGATGT	CAGCCTGGGG	GACGAGCTCC	ACTTAGACGG	CGAGGACGTG	GCGATGGCGC	4320
ATGCCGACGC	GCTAGACGAT	TTCGATCTGG	ACATGTTGGG	GGACGGGGAT	TCCCCGGGTC	4380
CGGGATTAC	CCCCACGAC	TCCGCCCCCT	ACGGCGCTCT	GGATATGGCC	GACTTCGAGT	4440
TTGAGCAGAT	GTTTACCGAT	GCCCTTGGA	TTGACGAGTA	CGGTGGGTAG	GGGGCGCGAG	4500
GATCCTCTAG	AGGGCCCTAT	TCTATAGTGT	CACCTAAATG	CTAGAGGATC	TTTGTGAAGG	4560
AACCTTACTT	CTGTGGTGTG	ACATAATTGG	ACAACTACC	TACAGAGATT	TAAAGCTCTA	4620
AGGTAAATAT	AAAATTTT	AGTGTATAAT	GTGTTAAACT	ACTGATTCTA	ATTGTTTGTG	4680
TATTTTAGAT	TCCAACCTAT	GGAAGTGTG	AATGGGAGCA	GTGGTGGAA	GCCTTTAATG	4740
AGGAAAACCT	GTTTTGCTCA	GAAGAAATGC	CATCTAGTGA	TGATGAGGCT	ACTGCTGACT	4800
CTCAACATTC	TACTCCTCCA	AAAAAGAAGA	GAAAGGTAGA	AGACCCCAAG	GACTTTCCTT	4860
CAGAATTGCT	AAGTTTTTTG	AGTCATGCTG	TGTTTAGTAA	TAGAACTCTT	GCTTGCTTTG	4920
CTATTTACAC	CACAAAGGAA	AAAGCTGCAC	TGCTATACAA	GAAAATTATG	GAAAAATATT	4980
TGATGTATAG	TGCCTTGACT	AGAGATCATA	ATCAGCCATA	CCACATTTGT	AGAGGTTTTA	5040
CTTGCTTTAA	AAAACCTCCC	ACACCTCCCC	CTGAACCTGA	AACATAAAAT	GAATGCAATT	5100
GTTGTTGTTA	ACTTGTTTAT	TGCAGCTTAT	AATGTTTACA	AATAAAGCAA	TAGCATCACA	5160
AATTTACAA	ATAAAGCATT	TTTTTCTACTG	CATTCTAGTT	GTGGTTTGTC	CAAACCTCATC	5220
AATGTATCTT	ATCATGTCTG	GATCATCCCG	CCATGGGTAT	ACATCGATGC	GGCCGC	5276

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7664 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGTACCACTA	GTAAGCTTAG	ATCCACTGTC	TGGGATTATA	TCAGGACAAC	CGAAGCCTGG	60
AAAGTGTATT	AGGTAGAGCA	TTTTCTTCCA	CGTGTTTGGG	CACGTTTCCG	ACAGCTAGGA	120
CTCCAGCTCT	GTCTTTGTAT	GTTACAGACT	GTAAATCAAT	CGCAGGTGAA	ACTGTTTGGA	180
CAGTAGGTGG	GGATCAAAGA	CCCTCCGCCC	GTGAGACTCT	AGGCGCTTTC	CCCTGCCACC	240
AGCCTGTCTC	CAGAGATGCT	CTGGAAGGAG	GCGGGCCCCG	GCGGTCTTTC	TGCTCTTTAG	300
CGTGCGCGAC	GCGGCGGCGG	GGGCAGGGCT	GGAGCAGAGA	GCGCTGCAGT	GATAGAACTT	360
TCTGACCCCG	CTGCGCAGGG	CGGCAGGGTG	GCAGGGTGGC	AGGGTGGCGA	GCTAAGCCAG	420
AGCCGAACGC	TGGAGCTCTG	GGAGGAACAT	CGAAGTGTTT	GTATGTGGTC	TGAGATCGGC	480
CTGACTATAT	TTTTTTGTCC	TAAATTTGCA	AGCACACACC	CACAAAGCTG	CGGTCTTGAC	540
CGGTATTCTT	TATAGAGCGC	AATGGAGTGA	GCTGAGTGTC	TAAACGATTT	CCCTAATTCA	600
TCTGATAGCA	GAGGCGCTCT	CCTAATTGGC	GAAGAGCTGC	CTCATGTCCG	CAACTTTTTG	660
GCAGAGTGAA	TTCCACAGCT	TTGTGTGTGT	GTGTGGGGGG	GGGTGTAAGG	GGTGTCTAAA	720
ACTTTCGGTC	TCCTACTATT	CTGTATCTCG	ACCGGTTGGT	TTTACACCCC	GGCTCATCTC	780
ATCAACGCAA	ACACCCCCAC	TCTCCTATGG	ACCCAAGGAC	CTGACGTGGG	GGAAGGTGGA	840
CATTAGGAAT	GTCAGAAACC	TAGAGTCCAC	GCTCCTCCTC	TCCATCTTTC	CACGAGTTTG	900
GGAAACTTCT	TGGCTGCGAA	GACTTTGACC	CACATCTGCA	TTTCTCAGCC	CCAGCTTCCA	960
AAAGTGCTGC	AGGTTCGGGA	GGGGAGACCT	CAGTCCTCCT	TTGTGAGGCT	TGTTTGCGTT	1020
GAGGGATTGG	CAGCGATGGC	TTCCAGATGG	GCTGAAACCC	TGCCCCGATT	TATTTAAACT	1080
GGTTCCTCGT	GGAGAGCTGT	GAATCGGGCT	CTGTATGCGC	TCGAGAAAAG	CCCCATTTCAT	1140
GAGAGGCAAG	GCCCAGTGGG	TCCCCCGGAC	TCCCCGACCC	CCCTCTCCCA	CAATATATCC	1200

CCCCTCCCTG	TGCCCCGCTG	CCGCCACCTC	CCGGGCTCCG	GCCCCGCGCG	CAGCGGCGAC	1260
GAAGCAACAC	AGTTCCCCGA	AAGAGGTAGC	TTTTTAATTG	GCCAGCCACA	AAGAATCACT	1320
TATGCCGCAC	GGCGGTAACG	AGGGGAACCG	GATCGGGCGG	CCAGGATGCT	ATCTGTGTAG	1380
CCCTTTTCGT	GCCACAATTA	GGGTGGTGCT	GGCTTCCTCC	GACCGCACCT	AGGCGATCTG	1440
GTTACACTGT	TGGCTCCTTT	CTTGGGCAGT	CATTTAATCC	TACTTTTTTAC	TCTACGAATG	1500
TCTGTCTGAT	GGAGGGCTGT	GTCCGGAGCC	CCATCCACAA	AGAGTCAGCC	AGCAGCTCTC	1560
ACACCCGGCT	GGATCTCATA	TGGTGCACCT	TCAGTACAAT	CTGCTCTGAT	GCCGCATAGT	1620
TAAGCCAGCC	AAGCTAGCTT	GCGCAAGCTA	GCTTGCGATC	CGTAAAAATG	TGTGAGAGTT	1680
ACAAAATGTC	TTCCGGGCTA	AGATCCGACA	GCCATGGTCC	AAAGAAGACT	TCGGCACTGC	1740
AGACTTAAAA	CCAGCTTTCT	AGCAGAGGCA	GAAGGATCTA	GAGCCAAAGG	CAAAGACTTG	1800
AATAGGCTGG	GAAGATGCAA	GAATGGCATT	TTACATAAAG	AACACTCTCT	CCTTTTCCAG	1860
CCAGCACACT	TGCATAGAAA	TTAAGTTTTA	CACCTGAAGT	TCTTTGTTTC	CATCCTGAGA	1920
AGCTCCAAAG	TCTGAGGTGG	TGTGGTATGC	TGGGTAAATC	TCCCCACCCC	CCAACATTCC	1980
CTGGGGGTTC	CATGGGGGTA	GCTTCTCCCA	AGGACTTCCA	GCGGCAACAC	AGAAATCCCA	2040
CTTCGAGACA	AAGGAGTTAC	TGCTTAAATC	AGGCCCTAAT	TTCCAAGGTT	CCCTTTGCTT	2100
AAAGTTCCTT	AGAGGACCAT	CTCACTTCTA	AAGAAAAGGT	GTATTCGGGG	ACCCATCCTC	2160
AACCTCCTTG	TTATGGAAGG	AGACTTCGGG	AACAGAGCAA	GGGCTGAGCC	TCCGGCAGTT	2220
TGGGGTAAGG	TTGGGGTTGG	GGGGAGCAAG	GAAGGCAAGT	GAGGCTGGAG	GCCCAGGGAT	2280
AGGGGAAGAT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	CTCGGGGATG	2340
GTGGTGGTGG	ACAAC TAGGA	AACTCTGGCG	CTTTCTCCTC	CCCTCACAAA	ACTGAGTCCA	2400
GCTGGAGCCG	CCTCCAGACT	CTCTGGCCAG	GGCCTCAGAG	TGGTCAACAG	TCCCTGGCCA	2460
GCGTTGCTCT	CTCCAGGCTA	AGGGCACCCA	CTCCCCTGGA	GATTCTTGAA	CCTGGGCCAG	2520
GAAGAGCCGA	ATTAGACAAG	TGTCTCCAAT	CCGGCTGCGT	GCGGATTTTG	TTGCGGTGTC	2580
GCTCGGTTGT	CTGCAGTTCC	TTTAGTCCCT	TCCCTGGCCT	GCCCCTTACA	CCTCCACACA	2640
GGTCCCCCTC	TGTGTAGGAA	TACACCAGAC	CCTCTCTTAG	CCACACACAC	CTCCAGTCCC	2700
GCGTCTACCT	AGATTTTTTTT	CATAGCTAGT	TGGATGGGGG	ATGGGTTAGG	GAGGCTGGGT	2760
GTGCGAGCCT	CCAGGTGGGA	GTTACCGGAC	AGGTACTCCG	CAAAGGAGCT	GGAAGGCAGG	2820
GCTGGAAAAC	TGTCCCCCAG	ATTTAGGATT	CTGGGCAGCT	TCCATCAGCT	TATACTTTGG	2880
CTCCCCCGCC	CCCTAAACTC	CCCATCCCCA	CCTTCCTTTC	TCCCGTTACT	TCGTCTCTCC	2940
TCGCCTTTCC	AGCCTTGAGT	CTAAAGCTCC	ATGCTTATGC	CTCTGCAAAC	AACCCCTCC	3000
CTTCTAACCC	CAGCAGAACT	CCGAGGAAAG	GGGCCGGAGG	CCCCCTTCT	CGCCTGTGGT	3060
TAGAGGGGGC	AGTGTGGCAG	TCCCAAGTGG	GGGCGACCGG	AGGCCGTCTC	GGTGCCCCGC	3120
ECGATCAGGC	CACTGGGCAC	ATCGGGGGCG	GGAAGCTGGG	CTCACCAAAG	GGGCGACTGG	3180
GCTTGGCAGG	TGTGGGCTCT	GGTCCGGCCT	GGGCAGGCTC	CGGGGGCGGG	GTCTCAGGTT	3240
ACAGCCCCGC	GGGGGGCTGG	GGGGCGGCCC	GCGGTTTGCG	CTGGTTTGCC	AGCCTTTGGA	3300
GCGACCGGGA	GCAATAAACC	GGAGCCTCTG	CTGGGAGAAG	ACGCAGAGCG	CCGCTGGGCT	3360
GCCGGGTCTC	CTGCCTCCTC	CTCCTGCTCC	TAGAGCCTCC	TGCATGAGGG	CGCGGTAGAG	3420
ACCCGGACCC	GCTCCGTGCT	CTGCCGCCTC	GCCGAGCTTC	GCCCCGAAGC	TGGGGAATTC	3480
GGATCCCCGG	GATCGAAAGA	GCCTGCTAAA	GCAAAAAAGA	AGTCACCATG	TCGTTTACTT	3540
TGACCAACAA	GAACGTGATT	TTCGTTGCCG	GTCTGGGAGG	CATTGGTCTG	GACACCAGCA	3600
AGGAGCTGCT	CAAGCGCGAT	CCCGTCGTTT	TACAACGTCT	TGACTGGGAA	AACCCTGGCG	3660
TTACCCAACT	TAATCGCCTT	GCAGCACATC	CCCCTTTCGC	CAGCTGGCTT	TATAGCGAAG	3720
AGGCCCCGAC	CGATCGCCCT	TCCCAACAGT	TGCGCAGCCT	GAATGGCGAA	TGGCGCTTTG	3780
CCTGGTTTCC	GGCACCAGAA	GCGGTGCCGG	AAAGCTGGCT	GGAGTGCGAT	CTTCCTGAGG	3840
CCGATACTGT	CGTCGTCCCC	TCAAACCTGGC	AGATGCACGG	TTACGATGCG	CCCATCTACA	3900
CCAACGTAAC	CTATTCCATT	ACGGTCAATC	CGCCGTTTGT	TCCCACGGAG	AATCCGACGG	3960
GTTGTTACTC	GCTCACATTT	AATGTTGATG	AAAGCTGGCT	ACAGGAAGGC	CAGACGCGAA	4020
TTATTTTTGA	TGGCGTTAAC	TTGGCGTTTC	ATCTGTGGTG	CAACGTGCGC	TGGGTGCGTT	4080
ACGGCCAGGA	CAGTCGTTTG	CCGTCTGAAT	TTGACCTGAG	CGCATTTTTA	CGCGCCGGAG	4140
AAAACCGCCT	CGCGGTGATG	GTGCTGCGTT	GGAGTGACGG	CAGTTATCTG	GAAGATCAGG	4200
ATATGTGGCG	GATGAGCGGC	ATTTTCCGTT	ACGTCTCGTT	GCTGCATAAA	CCGACTACAC	4260

AAATCAGCGA	TTTCCATGTT	GCCACTCGCT	TTAATGATGA	TTTCAGCCGC	GCTGAACTGG	4320
AGGCTGAAGT	TCAGATGTGC	GGCGAGTTGC	GTGACTACCT	ACGGGTAAACA	GTTTCTTTAT	4380
GGCAGGGTGA	AACGCAGGTC	GCCAGCGGCA	CCGCGCCTTT	CGGCGGTGAA	ATTATCGATG	4440
AGCGTGGTGG	TTATGCCGAT	CGCGTCACAC	TACGTCTGAA	CGTCGAAAAC	CCGAAACTGT	4500
GGAGCGCCGA	AATCCCGAAT	CTCTATCGTG	CGGTGGTTGA	ACTGCACACC	GCCGACGGCA	4560
CGCTGATTGA	AGCAGAAGCC	TGCGATGTCG	GTTTCCGCGA	GGTGCGGATT	GAAAATGGTC	4620
TGCTGCTGCT	GAACGGCAAG	CCGTTGCTGA	TTCGAGGCGT	TAACCGTCAC	GAGCATCATC	4680
CTCTGCATGG	TCAGGTCATG	GATGAGCAGA	CGATGGTGCA	GGATATCCTG	CTGATGAAGC	4740
AGAACAACCT	TAACGCCGTG	CGCTGTTTCG	ATTATCCGAA	CCATCCGCTG	TGGTACACGC	4800
TGTGCGACCG	CTACGGCCTG	TATGTGGTGG	ATGAAGCCAA	TATTGAAACC	CACGGCATGG	4860
TGCCAATGAA	TCTGCTGACC	GATGATCCGC	GCTGGCTACC	GGCGATGAGC	GAACGCGTAA	4920
CGCGAATGGT	GCAGCGCGAT	CGTAATCACC	CGAGTGTGAT	CATCTGGTCG	CTGGGGAATG	4980
AATCAGGCCA	CGGCGCTAAT	CACGACGCGC	TGTATCGCTG	GATCAAATCT	GTCGATCCTT	5040
CCCGCCCCGT	GCAGTATGAA	GGCGGCGGAG	CCGACACCAC	GGCCACCGAT	ATTATTTGCC	5100
CGATGTACGC	GCGCGTGGAT	GAAGACCAGC	CCTTCCCGGC	TGTGCCGAAA	TGGTCCATCA	5160
AAAAATGGCT	TTCGCTACCT	GGAGAGACGC	GCCCCTGAT	CCTTTGCGAA	TACGCCACCG	5220
CGATGGGTAA	CAGTCTTGCC	GGTTTCGCTA	AATACTGGCA	GGCGTTTCGT	CAGTATCCCC	5280
GTTTACAGGG	CGGCTTCGTC	TGGGACTGGG	TGGATCAGTC	GCTGATTAAT	TATGATGAAA	5340
ACGGCAACCC	GTGGTTCGGCT	TACGGCGGTG	ATTTTGGCGA	TACGCCGAAC	CATCGCCAGT	5400
TCTGTATGAA	CGGTCTGGTC	TTTGCCGACC	GCACGCCGCA	TCCAGCGCTG	ACGGAAGCAA	5460
AACACCAGCA	GCAGTTTTTC	CAGTTCGGTT	TATCCGGGCA	AACCATCGAA	GTGACCAGCG	5520
AATACCTGTT	CCGTCATAGC	GATAACGAGC	TCCTGCACTG	GATGGTGGCG	CTGGATGGTA	5580
AGCCGCTGGC	AAGCGGTGAA	GTGCCTCTGG	ATGTCGCTCC	ACAAGGTAAT	CAGTTGATTG	5640
AACTGCCTGA	ACTACCGCAG	CCGGAGAGCG	CCGGGCAACT	CTGGCTCACA	GTACGCGTAG	5700
TGCAACCGAA	CGCGACCGGA	TGGTCAGAAG	CCGGGCACAT	CAGCGCCTGG	CAGCAGTGGC	5760
GTCTGGCGGA	AAACCTCAGT	GTGACGCTCC	CCGCCGCGTC	CCACGCCATC	CCGCATCTGA	5820
CCACCAGCGA	AATGGATTTT	TGCATCGAGC	TGGGTAAATA	GCGTTGGCAA	TTTAACCGCC	5880
AGTCAGGCTT	TCTTTCACAG	CTGTGGATTG	GCGATAAAAA	ACAACCTGCT	ACGCCGCTGC	5940
GCGATCAGTT	CACCCGTGCA	CCGCTGGATA	ACGACATTGG	CGTAAGTGAA	GCGACCCGCA	6000
TTGACCTTAA	CGCCTGGGTC	GAACGCTGGA	AGGCGGCGGG	CCATTACCAG	GCCGAAGCAG	6060
CGTTGTTGCA	GTGCACGGCA	GATACACTTG	CTGATGCGGT	GCTGATTACG	ACCGCTCACG	6120
CGTGGCAGCA	TCAGGGGAAA	ACCTTATTTA	TCAGCCGGAA	AACCTACCGG	ATTGATGGTA	6180
GTGGTCAAAT	GGCGATTACC	GTTGATGTTG	AAGTGGCGAG	CGATACACCG	CATCCGGCGC	6240
CGATTGGCCT	GAAGTCCAG	CTGGCGCAGG	TAGCAGAGCG	GGTAAACTGG	CTCGGATTAG	6300
CGCCGCAAGA	AAACTATCCC	GACCGCCTTA	CTGCCGCCTG	TTTTGACCGC	TGGGATCTGC	6360
CATTGTGAGA	CATGTATACC	CCGTACGTCT	TCCCGAGCGA	AAACGGTCTG	CGCTGCGGGA	6420
CGCGCGAATT	GAATTATGGC	CCACACCAGT	GGCGCGGCGA	CTTCCAGTTC	AACATCAGCC	6480
GCTACAGTCA	ACAGCAACTG	ATGGAAACCA	GCCATCGCCA	TCTGCTGCAC	GCGGAAGAAG	6540
GCACATGGCT	GAATATCGAC	GGTTTCCATA	TGGGGATTGG	TGGCGACGAC	TCCTGGAGCC	6600
CGTCAGTATC	GGCGGAATTA	CAGCTGAGCG	CCGGTCGCTA	CCATTACCAG	TTGGTCTGGT	6660
GTCAAAAATA	ATAATAACCG	GCAGGCCATG	TCTGAAAGTA	TTGCGGTAAG	GAAATCCATT	6720
ATGTACTATT	TAAAAACAC	AAACTTTTGG	ATGTTTCGGT	TATTCTTTT	CTTTTACTTT	6780
TTTATCATGG	GAGCCTACTT	CCCGTTTTTC	CCGATTTGGC	TACATGACAT	CAACCATATG	6840
AGCAAAAGTG	ATACGGGTAT	TATTTTTGCC	GCTATTTCTC	TGTTGTCTGCT	ATTATTCCAA	6900
CCGCTGTTGG	TCTGCTTTCT	GACAAACTCG	GCCTCGACTC	TAGACTGAGA	ACTTCAGGGT	6960
GAGTTTGGGG	ACCCTTGATT	GTTCTTTCTT	TTTCGCTATT	GAAAAATTCA	TGTTATATGG	7020
AGGGGGCAAA	GTTTTTCAGGG	TGTTGTTTAG	AATGGGAAGA	TGTCCCTTGT	ATCACCATGG	7080
ACCCTCATGA	TAATTTTGTG	TCTTTCACTT	TCTACTCTGT	TGACAACCAT	TGTCTCCTCT	7140
TATTTTCTTT	TCATTTTCTG	TAACTTTTTT	CGTTAAACTT	TAGCTTGCA	TTGTAACGAA	7200
TTTTTAAATT	CACTTTCGTT	TATTTGTCAG	ATTGTAAGTA	CTTTCTCTAA	TCACTTTTTT	7260
TTCAAGGCAA	TCAGGGTAAT	TATATTGTAC	TTCAGCACAG	TTTAGAGAA	CAATTGTTAT	7320

AATTAAATGA	TAAGGTAGAA	TATTTCTGCA	TATAAATTCT	GGCTGGCGTG	GAAATATTCT	7380
TATTGGTAGA	AACAAC TACA	TCCTGGTAAT	CATCCTGCCT	TTCTCTTTAT	GGTTACAATG	7440
ATATACTG	TTTGAGATGA	GGATAAAATA	CTCTGAGTCC	AAACCGGGCC	CCTCTGCTAA	7500
CCATGTTTAT	GCCTTCTTCT	TTTTCTTACA	GCTCCTGGGC	AACGTGCTGG	TTGTTGTGCT	7560
GTCTCATCAT	TTTGGCAAAG	AATTCAC TCC	TCAGGTGCAG	GCTGCCTATC	AGAAGGTGGT	7620
GGCTGGTGTG	GCCAATGCCC	TGGCTCACAA	ATACCACTGA	GATC		7664

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGAGGGCCTG CTCGATCTCC

20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCATTCCAC CACTGCTCCC

20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCACCTT CTCATGACCT C

21

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GTTGGTGTAG ATGGGCGCAT CG

22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGGGGTCTC AGGTTACAGC C

21

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCCTCTGGC CTGCTGGCTC ATG

23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CAGGAGAGTC TTGCCTGTAT CCTC

24

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1521 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CAAGATGCAT	CCAGGGGTCC	TGGCTGCCTT	CCTCTTCTTG	AGCTGGACTC	ATTGTCGGGC	60
CCTGCCCCTT	CCCAGTGGTG	GTGATGAAGA	TGATTTGTCT	GAGGAAGACC	TCCAGTTTGC	120
AGAGCGCTAC	CTGAGATCAT	ACTACCATCC	TACAAATCTC	GCGGGAATCC	TGAAGGAGAA	180
TGCAGCAAGC	TCCATGACTG	AGAGGCTCCG	AGAAATGCAG	TCTTTCTTCG	GCTTAGAGGT	240
GACTGGCAAA	CTTGACGATA	ACACCTTAGA	TGTCATGAAA	AAGCCAAGAT	GCGGGGTTGT	300
CGATGTGGG	TGAATACAATG	TTTTCCCTCG	AACTCTTAAA	TGGTCCAAAA	TGAATTTAAC	360
CTACAGAATT	GTGAATTACA	CCCCTGATAT	GACTCATTCT	GAAGTCGAAA	AGGCATTCAA	420
AAAAGCCTTC	AAAGTTTGGT	CCGATGTAAC	TCCTCTGAAT	TTTACCAGAC	TTCACGATGG	480
CATTGCTGAC	ATCATGATCT	CTTTTGGAAT	TAAGGAGCAT	GGCGACTTCT	ACCCATTTGA	540
TGGGCCCTCT	GGCCTGCTGG	CTCATGCTTT	TCCTCCTGGG	CCAAATTATG	GAGGAGATGC	600
CCATTTTGAT	GATGATGAAA	CCTGGACAAG	TAGTTCCAAA	GGCTACAAC	TGTTTCTTGT	660
TGCTGCGCAT	GAGTTCGGCC	ACTCCTTAGG	TCTTGACCAC	TCCAAGGACC	CTGGAGCACT	720
EATGTTTCCT	ATCTACACCT	ACACCGGCAA	AAGCCACTTT	ATGCTTCCTG	ATGACGATGT	780
ACAAGGGATC	CAGTCTCTCT	ATGGTCCAGG	AGATGAAGAC	CCCAACCCTA	AACATCCAAA	840
AACGCCAGAC	AAATGTGACC	CTTCCTTATC	CCTTGATGCC	ATTACCAGTC	TCCGAGGAGA	900
AACAATGATC	TTTAAAGACA	GATTCTTCTG	GCGCCTGCAT	CCTCAGCAGG	TTGATGCGGA	960
GCTGTTTTTA	ACGAAATCAT	TTTGGCCAGA	ACTTCCCAAC	CGTATTGATG	CTGCATATGA	1020
GCACCCTTCT	CATGACCTCA	TCTTCATCTT	CAGAGGTAGA	AAATTTTGGG	CTCTTAATGG	1080
TTATGACATT	CTGGAAGGTT	ATCCCAAAAA	AATATCTGAA	CTGGGTCTTC	CAAAAGAAGT	1140
TAAGAAGATA	AGTGCAGCTG	TTCACCTTGA	GGATACAGGC	AAGACTCTCC	TGTTCTCAGG	1200
AAACCAGGTC	TGGAGATATG	ATGATACTAA	CCATATTATG	GATAAAGACT	ATCCGAGACT	1260
AATAGAAGAA	GACTTCCCAG	GAATTGGTGA	TAAAGTAGAT	GCTGTCTATG	AGAAAAATGG	1320
TTATATCTAT	TTTTTCAACG	GACCCATACA	GTTTGAATAC	AGCATCTGGA	GTAACCGTAT	1380
TGTTTCGCGTC	ATGCCAGCAA	ATTCCATTTT	GTGGTGTTAA	GTGTCTTTTT	AAAAATTGTT	1440
ATTTAAATCC	TGAAGAGCAT	TTGGGGTAAT	ACTTCCAGAA	GTGCGGGGTA	GGGGAAGAAG	1500
AGCTATCAGG	AGAAAGCTTG	G				1521

(2) INFORMATION FOR SEQ ID NO:19

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Pro Arg Cys Gly Xaa Pro Asp

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

His Glu Xaa Gly His Xaa Xaa Xaa Xaa Xaa His Ser  
 1 5 10

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 471 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met His Pro Gly Val Leu Ala Ala Phe Leu Phe Leu Ser Trp Thr His  
 1 5 10 15  
 Cys Arg Ala Leu Pro Leu Pro Ser Gly Gly Asp Glu Asp Asp Leu Ser  
 20 25 30  
 Glu Glu Asp Leu Gln Phe Ala Glu Arg Tyr Leu Arg Ser Tyr Tyr His  
 35 40 45  
 Pro Thr Asn Leu Ala Gly Ile Leu Lys Glu Asn Ala Ala Ser Ser Met  
 50 55 60  
 Thr Glu Arg Leu Arg Glu Met Gln Ser Phe Phe Gly Leu Glu Val Thr  
 65 70 75 80  
 Gly Lys Leu Asp Asp Asn Thr Leu Asp Val Met Lys Lys Pro Arg Cys  
 85 90 95  
 Gly Gly Val Asp Val Gly Glu Tyr Asn Val Phe Pro Arg Thr Leu Lys  
 100 105 110  
 Trp Ser Lys Met Asn Leu Thr Tyr Arg Ile Val Asn Tyr Thr Pro Asp  
 115 120 125  
 Met Thr His Ser Glu Val Glu Lys Ala Phe Lys Lys Ala Phe Lys Val  
 130 135 140  
 Trp Ser Asp Val Thr Pro Leu Asn Phe Thr Arg Leu His Asp Gly Ile  
 145 150 155 160  
 Ala Asp Ile Met Ile Ser Phe Gly Ile Lys Glu His Gly Asp Phe Tyr

[illegible]